



Center for
Environmental
Information and
Statistics

US Environmental
Protection Agency

Major Findings from the CEIS Review of EPA'S

SAFE DRINKING WATER INFORMATION SYSTEM (SDWIS) DATABASE



April 2, 1999

Major Findings from the CEIS Review of EPA's SDWIS Database

1. INTRODUCTION	1
2. SUMMARY ANSWERS TO REVIEW QUESTIONS	2
3. DETAILED ANSWERS TO REVIEW QUESTIONS	4
3.1. What does the database cover?	4
3.2. Can the database be used for spatial analysis?	6
3.3. Can the database be used for temporal analysis?	6
3.4. How consistent are the variables over space and time?	6
3.5. Can data from SDWIS be linked with information from other databases?	6
3.6. How accurate are the data in SDWIS?	7
3.7. What are the limitations of SDWIS?	7
3.8. How can I get information on SDWIS?	7
3.9. Is there documentation on SDWIS?	7

1. INTRODUCTION

The **Safe Drinking Water Information System/Federal version (SDWIS)** is an EPA national database storing information about the nation's drinking water. Designed to replace the system known as the Federal Reporting Data System (FRDS), SDWIS stores the information EPA needs to monitor public water systems (PWS). The database covered over 172,000 PWSs in 1996, the most recent year for which annual figures have been compiled. SDWIS tracks information on drinking water contamination levels as required by the 1974 Safe Drinking Water Act and its 1986 and 1996 amendments.

SDWIS is operated and maintained by the Office of Ground Water and Drinking Water within the Office of Water. This office uses the information reported to SDWIS to oversee State drinking water programs, track contaminant levels, respond to public inquiries, and prepare national reports. EPA also uses this information to evaluate the effectiveness of its programs and regulations, and to determine whether new regulations are needed to further protect public health.

SDWIS is one of the major EPA databases that are being reviewed to characterize their overall quality and applicability.

2. SUMMARY ANSWERS TO REVIEW QUESTIONS

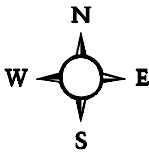


2.1. What does the database cover?

The SDWIS database contains records for over 172,000 public water system (PWS) facilities that serve over 90% of the US population (1996 data). PWS facilities that are required to report water quality data:

- i. provide piped water for human consumption to at least 15 service connections; or
- ii. serve an average of at least 25 people for at least 60 days.

These facilities report data on monitored contaminant levels for EPA regulated contaminants as well as other contaminants. As of 1996, EPA regulations establish Maximum Contaminant Levels (MCLs) for 72 individual contaminants and Treatment Techniques (TTs) for nine individual contaminants. These contaminants include microbiological contaminants, radionuclides, inorganic chemicals, and organic chemicals.



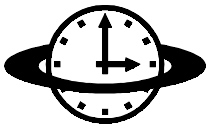
2.2. Can the database be used for spatial analysis?

Yes. Currently, spatial information contained within the database includes PWS addresses, hydrologic unit codes (HUC), and latitude/longitude information (including those generated from ZIP code centroids). SDWIS reporting rules will require all PWS to report latitude and longitude information for water intake source by the year 2004.



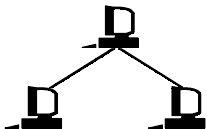
2.3. Can the database be used for temporal analysis?

Yes. There is temporal information in SDWIS as the violation dates are available. Monitoring data are available only when monitoring results exceed the MCL. Trends analysis may be performed as the dataset is frozen every quarter.



2.4. How consistent are the variables over space and time?

EPA-promulgated rules dictate what and how often contaminants are monitored. These rules are subject to modification as environmental situations demand. Therefore, the consistency of variables over space and time is not assured.



2.5. Can data from SDWIS be linked with information from other databases?

Yes. SDWIS data have the potential to be linked with data from other databases using latitude/longitude, identifiers for city and State, or hydrologic unit code (HUC).



2.6. How accurate are the data in SDWIS?

The data in SDWIS are becoming more accurate as the system matures, as reporting bugs are removed, and as the data become open to greater public scrutiny and review. According to the Program Office, the gross error rate in the information for determining whether systems are in significant noncompliance (with drinking water standards) is roughly 10-12%.



2.7. What are the limitations of SDWIS?

Reliance on PWS' monitoring and reporting leads to uneven data due to under reporting. Since states are only required to report violations, EPA can only know the extent of under-reporting by auditing state records. States and Regions are responsible for data collection. There are differences among the States in the emphasis placed on data collection efforts and responsiveness to notifications of data errors. In addition, the nature of the Safe Drinking Water Act allows contaminant monitoring programs to be modified, limiting the comparability of data over time and across public water systems.



2.8. How can I get information on SDWIS?

A subset of SDWIS is available to the public through EPA's *Envirofacts* website, http://www.epa.gov/enviro/html/sdwis/sdwis_ov.html. In addition, FOIA requests may be submitted to obtain SDWIS data.



2.9. Is there documentation on SDWIS?

There are a number of user and system documents available from the Office of Drinking Water and Ground Water that are listed in the detailed answers that follow.

3. DETAILED ANSWERS TO REVIEW QUESTIONS



Regulated Contaminants in 1996

- Arsenic
- Copper*
- Fluoride
- Lead*
- Total Coliforms
- Total Trihalomethanes (TTHM)
- Benzene
- Carbon Tetrachloride
- 1,2-Dichloroethane
- 1,1-Dichloroethylene
- *p*-Dichlorobenzene
- 1,1,1-Trichloroethane
- Trichloroethylene
- Vinyl Chloride
- Acrylamide*
- Alachlor (Lasso)
- Asbestos
- Atrazine
- Barium
- Carbofuran
- Cadmium
- Chlordane
- Chromium
- Dibromochloropropane (DBCP)
- *o*-Dichlorobenzene
- *cis*-1,2-Dichloroethylene
- *trans*-1,2-Dichloroethylene
- 1,2-Dichloropropane
- 2,4-D
- 2,4,5-TP (Silvex)
- Ethylbenzene
- Ethylene Dibromide (EDB)
- Epichlorohydrin*
- Heptachlor
- Heptachlor Epoxide
- Lindane (BHC-gamma)
- Mercury
- Methoxychlor
- Monochlorobenzene
- Nitrate
- Nitrite
- PCBs
- Pentachlorophenol
- Selenium
- Styrene
- Tetrachloroethylene
- Toluene

3.1. What does the database cover?

Under the Safe Drinking Water Act (SDWA), EPA is responsible for regulating the water provided by Public Water Systems. Some of these systems use surface water from a river or lake and serve millions of urban customers. Others use wells or ground water and serve only a few dozen visitors to a campground or gas station. A Public Water System (PWS) provides piped water for human consumption to at least 15 service connections or serves an average of at least 25 people for at least 60 days each year. The water systems included in SDWIS provide water to over 90% of the US population. These PWSs can be community, nontransient noncommunity, or transient noncommunity systems and are defined as follows.

- A Community Water System (CWS) is a PWS that provides water to the same population year round.
- A Nontransient Noncommunity Water System (NTNCWS) is a PWS that regularly serves at least 25 of the same people at least six months of the year. Examples of these systems include schools, factories, and hospitals that have their own water supplies.
- A Transient Noncommunity Water System (TNCWS) caters to transitory customers and include campgrounds, restaurants, motels, and gas stations.

The Safe Drinking Water Information System (SDWIS) contains information about these water systems. SDWIS contained records for over 172,000 facilities in 1996. As of 1996, EPA regulations establish Maximum Contaminant Levels (MCLs) for 72 individual contaminants and Treatment Techniques (TTs) for nine individual contaminants. This list of 81 is comprised of 6 microbiological contaminants, 4 radionuclides, 17 inorganic chemicals, and 54 organic chemicals.

Who Must Report?

All Public Water Systems must report water quality information for inclusion in the SDWIS database. These systems have different monitoring and reporting requirements based on the population served. The PWS size categories are as follows.

<u>System Size</u>	<u>Population Served</u>
Very Small	25-500
Small	501-3,300
Medium	3,301-10,000
Large	10,001-100,000
Very Large	> 100,000

- Toxaphene
- Xylenes (total)
- Antimony
- Beryllium
- Cyanide
- Dalapon
- Di(2-ethylhexyl)adipate
- Di(2ethylhexyl) phthalate
- Dichloromethane
- Dinoseb
- Dioxin (2,3,7,8-TCDD)
- Diquat
- Endothall
- Endrin
- Glyphosate
- Hexachlorobenzene (HCB)
- Hexachlorocyclopentadiene
- Oxamyl (Vydate)
- PAHs (Benzo(a)pyrene)
- Picloram
- Simazine
- Thallium
- 1,2,4-Trichlorobenzene
- 1,1,2-Trichloroethane
- Beta Particle and Photon Radioactivity
- Gross Alpha Particle Activity
- Radium-226
- Radium-228
- *Giardia lamblia**
- *Legionella**
- Heterotrophic Plate Count*
- Turbidity*
- Viruses*

* Denotes contaminants which are regulated by treatment techniques (TTs) instead of a maximum contaminant level (MCL).

How are data reported?

Regulated Public Water Systems are responsible for monitoring their own systems and collecting and reporting compliance with water quality standards. Collected data are forwarded to a Primacy Agent. Primacy agents are usually State or Tribal-level departments of environmental protection (or similarly titled agencies), although this role may fall to EPA Regional offices if State or Tribal governments choose not to apply for primacy. EPA recognition of primacy is given on the condition that primacy agents develop and enforce State or Tribal-government water quality regulations.

There is a reporting schedule that varies by type of information. Data on water system, water system facility, legal entity, geographic area, service area, visits, sources, and treatments are required to be reported annually, although the primacy states are encouraged to submit them quarterly. The annual inventory data are usually submitted between September and December as this data is used to calculate the grant formula in January and February. Enforcement, violations, sample, Lead and Copper Rule (LCR) 90th percentile data, and milestone data are reported on a quarterly basis. Additionally, long term compliance monitoring periods vary and a system is not out of compliance until the period ends. For example, with a three year monitoring period, the system is not out of compliance until the three years end when States would report the monitoring violation 45 days after the end of the third year. The violation would be described as a violation that covered the three year period. Currently, primacy States and Regional offices (in the case of Wyoming, District of Columbia, and all Indian lands) submit data to SDWIS electronically. EPA continually updates SDWIS data and the dataset is frozen quarterly, for analytical purposes.

Data Elements

SDWIS contains an inventory of public water systems with hundreds of variables. The following are a selection of data elements. Many of the elements listed below are not required to be reported and are therefore not available for all public water systems.

facility - name, identification number, Facility Index System (FINDS) number, activity status, latitude and longitude, HUC, sampling point

water system - name, identification number, activity status, city served, source of water, population count

legal entity - contacts, street address, city, county, zip code, employer identification number

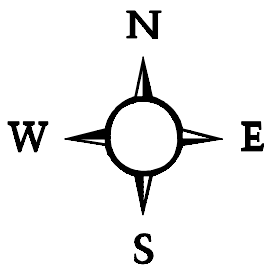
contaminants - SDWIS code, Chemical Abstract Service (CAS) number, MCL measure, common name, monitoring data when sample result exceeds the MCL

treatment -objective, process type

geographic area served - name, Federal Information Processing Standards (FIPS) code, State defined code

water purchase - type, quantity, source

flow - quantity, connection and disconnection date, connection type



3.2. Can the database be used for spatial analysis?

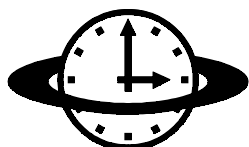
Yes. Data, when available, may be analyzed at numerous spatial levels, including national, Regional, State, county, community, and ZIP code. Records have facility specific latitude and longitude values, including those generated from ZIP codes. In accordance with EPA policy, latitude and longitude will have to be reported for each PWS water intake point on a phased basis, beginning in 2000 and complete by 2004.



3.3. Can the database be used for temporal analysis?

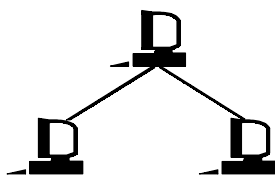
Yes. SDWIS data are updated continuously as they are received from States, Regional offices, and other primacy agents. Data are available for all violations and when systems return to compliance. It is difficult to analyze trends in contaminants since only levels exceeding the MCL are reported.

There is a standard reporting timeframe. Inventory is required annually, all other data is required quarterly. Any violations that are identified within the last quarter must be reported within 45 days of the end of the quarter within which they were identified. For instance, some rules specify the reporting of annual averages. Thus, one may not know of a violation until after the close of the fourth quarter. Reporting requirements and schedules therefore can have substantial impact on the ability to do trends analyses.



3.4. How consistent are the variables over space and time?

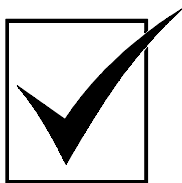
Rules promulgated by EPA under the Safe Drinking Water Act dictate collection frequencies as well as the list of contaminants to be monitored. Both the list of contaminants and their collection frequencies are subject to modification. As mentioned above, collection frequencies and protocols may vary with the type and size of PWS. In addition, EPA is able to require the monitoring of additional contaminants should the need arise. Thus, the universe of variables is subject to change, limiting the database's consistency.



3.5. Can data from SDWIS be linked with information from other databases?

Yes. Data from SDWIS have the potential to be linked with data from other databases using some of the variables listed below.

- name of facility and facility ID
- latitude and longitude
- codes for cities, counties, and States
- HUC codes
- CAS numbers



3.6. How accurate are the data in SDWIS?

The Safe Drinking Water Act (SDWA) allows States and territories to seek EPA approval to administer their own Public Water Supply Supervision (PWSS) programs. The authority to run a PWSS program is called primacy. Primacy States report quarterly to EPA on their Public Water System (PWS) inventory statistics, the incidence of violations, and the enforcement actions taken against violators. The accuracy of the information in the SDWIS database depends upon the quality of information provided by the primacy agents.

The gross error rate for significant noncompliance determinations is roughly 10-12% despite aggressive attempts to improve data collection and reporting procedures and regular quality assurance/quality control inspections by Regional offices. EPA provides funding to States and Regions for data quality improvement projects. Water systems, if they discover errors in their SDWIS entries, must go through the primacy agency to submit a correction. The relational integrity of SDWIS data is tested and evaluated periodically by the database administrator, and incoming data are screened for accuracy.



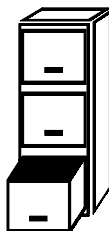
3.7. What are the limitations of SDWIS?

- The use of State and Tribal governments as primacy agents responsible for reporting and correcting SDWIS data results in substantial variation in data quality. While some are very responsive in correcting errors and reporting data regularly, others are less so.
- Reporting inconsistencies also result in an incomplete database.
- The addition of rule-driven data collection activities as well as the modification of existing activity parameters results in an ever-changing database. This limits its utility in performing temporal and spatial analyses of drinking water quality.



3.8. How can I get information on SDWIS?

SDWIS information may also be accessed through a Freedom of Information Act (FOIA) request or through *Envirofacts*. Through filing a FOIA request, individuals can access the information contained in SDWIS. These requests are processed through EPA's Office of Ground Water and Drinking Water. A subset of SDWIS information is available on the internet at EPA's *Envirofacts* website, http://www.epa.gov/enviro/html/sdwis/sdwis_ov.html.



3.9. Is there documentation on SDWIS?

Yes. The rules that dictate what contaminants are monitored, their monitoring schedules, and the manner in which they are monitored are printed in the Code of Federal Regulations (CFR). The Office of Drinking Water and Ground Water also provides numerous guidance documents for PWS monitoring personnel. These include, among others, guidance on the implementation of rules and use of the electronic reporting system. There is also a data element dictionary and the following documents.

PWS Inventory and Compliance Statistics: FY1991-FY1995

“System Users Guide for SDWIS/FED” Prepared by EPA Systems Development Center, Arlington, VA. November, 1997.